Cochrane Database of Systematic Reviews

Contrast-enhanced ultrasound compared with computed tomography, magnetic resonance imaging, and positron emission tomography for diagnosing liver metastases in people with newly diagnosed colorectal cancer (Protocol)

Lund M, Bjerre TA, Grønbæk H, Mortensen F, Kragh Andersen P

Lund M, Bjerre TA, Grønbæk H, Mortensen F, Kragh Andersen P.
Contrast-enhanced ultrasound compared with computed tomography, magnetic resonance imaging, and positron emission tomography for diagnosing liver metastases in people with newly diagnosed colorectal cancer.
Cochrane Database of Systematic Reviews 2016, Issue 10. Art. No.: CD012388.
DOI: 10.1002/14651858.CD012388.

www.cochranelibrary.com

TABLE OF CONTENTS

[Diagnostic Test Accuracy Protocol]

Contrast-enhanced ultrasound compared with computed tomography, magnetic resonance imaging, and positron emission tomography for diagnosing liver metastases in people with newly diagnosed colorectal cancer

Martin Lund[1], Thomas Abramovitz Bjerre[2], Henning Grønbæk[3], Frank Mortensen[4], Per Kragh Andersen[5]

[1]Department of Public Health Programmes, Randers Regional Hospital, Randers, Denmark. [2]Diagnostic Imaging Department, Randers Regional Hospital, Randers, Denmark. [3]Medical Department V, Aarhus University Hospital, Århus, Denmark. [4]Department of Surgery L, Aarhus University Hospital, Aarhus, Denmark. [5]Department of Biostatistics, University of Copenhagen, Faculty of Health Sciences, Institute of Public Health, Copenhagen K, Denmark

Contact address: Martin Lund, Department of Public Health Programmes, Randers Regional Hospital, Skovlyvej 15, Randers, 8930 NØ, Denmark. martlund@rm.dk.

Editorial group: Cochrane Hepato-Biliary Group.
Publication status and date: New, published in Issue 10, 2016.

Citation: Lund M, Bjerre TA, Grønbæk H, Mortensen F, Kragh Andersen P. Contrast-enhanced ultrasound compared with computed tomography, magnetic resonance imaging, and positron emission tomography for diagnosing liver metastases in people with newly diagnosed colorectal cancer. *Cochrane Database of Systematic Reviews* 2016, Issue 10. Art. No.: CD012388. DOI: 10.1002/14651858.CD012388.

ABSTRACT

This is the protocol for a review and there is no abstract. The objectives are as follows:

To compare the accuracy of CEUS, CECT, MRI, and 18F-FDG PET-CT for diagnosing liver metastases in people with newly-diagnosed colorectal cancer.

Potential sources of heterogeneity

We will investigate the following potential sources of heterogeneity:

- the use of different reference standards.
- different ways of selecting the study populations, e.g. different inclusion and exclusion criteria.
- different locations of the study populations (country, state, region).
- age of participants in the study population.
- sex of participants in the study population.
- differences in clinician skills for the performance of CEUS.

BACKGROUND

Colorectal cancer is one of the most common cancer diseases in the United States and in the rest of the Western world. The probability of developing colorectal cancer in a lifetime is for American men 4.8% (1 in 21) and 4.5% (1 in 22) for American women (Siegel 2015). From the time of diagnosis, the five-year relative survival rates in the United States, adjusted for normal life expectancy, are 65% for colon cancer and 68% for rectum cancer (Siegel 2015).

Since the mid-1980s, there has been a decline in colorectal cancer in the United States. Between 2008 and 2011, the decline has been 4% or greater per year. Most likely, the decline reflects the increased uptake of screening, primarily in the form of colonoscopy. Colonoscopy can prevent cancer by the removal of precancerous lesions (Siegel 2015). In the United States, the use of colonoscopies among adults aged 50 to 75 years increased from 19.1% in 2000 to 54.5% in 2013 (Siegel 2015).

Metastatic disease to the liver is a very common clinical situation in oncology, and the liver is the most common site of metastatic spread from colorectal cancer. With the diagnosis of colorectal cancer, liver metastases may be synchronous, i.e. diagnosed at the same time as the primary tumour, or metachronous, i.e. develop during follow-up after surgical resection of the primary tumour. Surgical resection, stereotactic radiation therapy, and radiofrequency ablation of liver metastases are options for curative treatment in colorectal cancer with the presence of a limited number of metastases (Cirocchi 2012). However, it is a clinical challenge to diagnose the presence and exact localization of liver metastases at an early state, when curative treatment is still an option.

Around 20% to 25% of all people with colorectal cancer have metastatic spread at the time of diagnosis, and approximately 50% of all these people develop liver metastases during the course of colorectal cancer (Kanas 2012; Vatandoust 2015). Metastases confined to the liver at the time of the detection of colorectal cancer are potentially resectable in about 10% to 30% of the patients (Kanas 2012; Vatandoust 2015). Hepatic resection is considered to be the best curative treatment for liver-limited colorectal metastases. However, there are some contraindications to hepatic resection which include unresectable extrahepatic disease, more than 70% metastatic liver involvement, liver failure, and being surgically unfit (Vatandoust 2015). The lungs are the second most common site of distant metastases in people with colorectal cancer, and the peritoneum is the third most common site (Vatandoust 2015). In people with isolated hepatic lesions, five-year survival after surgical resection is reported to range from 16% to 74% (median 38%) (Kanas 2012), and in another study five-year survival is reported to range from 25% to 58%, and 10-year survival is reported to range from 17% to 28% (Vatandoust 2015). It is therefore important to detect and treat colorectal liver metastases as early as possible in the development of the disease, and thus be able to offer the patients the best possible treatment. The first step in this process is a reliable triage test to detect the liver metastases, if any are present. The next step in the process of treating liver metastases depends on the imaging technique, to decide which patients may be surgical candidates. Thus, the ability of the imaging technique to demonstrate the exact number, the size, the regional distribution, and the volume of the remaining liver is crucial to determine resectability.

A meta-analysis published in 2010 assessed the accuracy of computed tomography (CT), magnetic resonance imaging (MRI), fluro-18-deoxyglucose (FDG) positron emission tomography (PET) or FDG PET-CT or both in the detection of colorectal liver metastases (Niekel 2010). The meta-analysis included prospective studies and treatment-naïve people. The authors concluded that MRI was the preferred first-line modality for evaluating colorectal liver metastases. FDG PET could be used as a second-line modality. The evidence on the role of FDG PET-CT was unclear, due to a small number of studies.

A Cochrane Review with the primary objective of determining the diagnostic accuracy of integrated FDG PET-CT as a replacement test for conventional imaging for the pre-operative staging of recurrent colorectal cancer is still in preparation (Crawford 2012). The comparisons of interest are PET-CT versus clinical follow-up including standard imaging as a replacement test for the detection of extra-hepatic and intra-hepatic lesions.

Target condition being diagnosed

The clinical target condition of this review is colorectal liver metastases in people with newly-diagnosed colorectal cancer.

Index test(s)

Contrast-enhanced ultrasound (CEUS)

An ultrasound scanner is a medical imaging modality based on the use of echoes from ultrasound waves to produce live pictures of all kinds of anatomical structures, like the musculoskeletal system, and the parenchyma of the inner organs. The ultrasound transducer is handheld and is placed directly on the skin in the anatomical area of interest. CEUS is basically the same as conventional ultrasound, but adds a contrast agent. The contrast agent is administered as an intravenous infusion and makes it possible to study liver perfusion in live pictures. It is possible to characterise focal liver lesions with patterns of enhancement, due to the use of the contrast agent. The contrast agent consists of microbubbles (sulphur hexafluoride) and is without any known serious adverse effects (Solbiati 2003). The advantages of CEUS are that it is performed without the use of ionizing radiation, and the examination is therefore considered to cause no harm to the human body (Solbiati 2003). It is a relatively fast examination which can be performed in approximately 30 minutes. If necessary, it is possible to perform a biopsy of suspected lesions in the liver during the examination. The disadvantage is that the value of CEUS depends very much on the skills of the physician who performs the scan (Solbiati 2003), and even if it is possible to store the images, there is no guarantee that other physicians will be able to interpret them. Ultrasound scanners are widely available at clinical centres in most countries, because the cost is relatively low.

Contrast-enhanced computed tomography (CECT)

A CT scanner is a medical imaging modality based on x-rays. The pictures are acquired while the person is moving through a circle-shaped gantry, where the x-ray tube and the chain of detectors are circling at high speed. A contrast agent is administered as an intravenous infusion by an automatic syringe CT injector to enhance the perfusion of the inner organs**.** The advantages of CECT are that the examination time is very short, lasting from five to 15 minutes, and the images are ready for interpretation immediately. CECT may also examine extra-hepatic tissue and organs. Similar to CEUS, it is possible to characterize focal liver lesions with patterns of enhancement, due to the use of the contrast agent. The images are stored electronically and are available to other physicians, and CT volumetry allows for volume estimation of the future liver remnant in the case of hepatic resection (Lim 2014). The disadvantages are the use of relatively high doses of ionising radiation to the patients and the use of contrast agents with iodine, which are known to have certain high-risk adverse effects, such as allergic reactions and, in the worst case, anaphylactic shock. One of the contraindications to CECT is therefore previous allergic reactions to contrast agents with iodine. Another contraindication is renal insufficiency, due to increased risk of kidney failure. CT scanners are widely available at clinical centres in many countries, even though the cost is relatively high (NCHS 2010).

Magnetic resonance imaging (MRI)

A MRI scanner is a medical imaging modality based on the use of a strong magnetic field and radio waves to produce images of the body. The magnet is in the form of a large tube, and the person is placed inside it. A MRI scanner is quite noisy, but new technology has made it possible to make silent protocols to some degree in the newest machines. The advantages of MRI are similar to CEUS. It is performed without the use of ionising radiation, and the examination is therefore considered to cause no harm to the human body (Westbrook 2011). There are two different kinds of intravenously-administered contrast agents, which can be used to detect liver metastases on MRI: gadolinium (Gd) and ferucarbotran (SuperParamagnetic Iron Oxide, (SPIO)). These contrast agents are not known to have serious adverse effects if they are used in small doses. However, people with renal insufficiency are known to be at high risk of adverse effects like nephrogenic systemic fibrosis if they are exposed to MRI gadolinium agents, especially if the gadolinium agents are used in high doses. Another disadvantage is that MRI is a time-consuming examination compared to CEUS and CECT. It takes approximately one hour to perform a MRI scan of the liver. Contraindications to a MRI examination are people suffering from claustrophobia because of the position inside the magnetic tube, and also people with metal implants, especially in the head or eyes, which are very likely to move due to the magnetic field. People with claustrophobia are often able to complete a MRI examination in an open MRI scanner. MRI scanners are widely available at clinical centres in the western world, but many countries in other parts of the world do not have extensive access to MRI scanners (NCHS 2010). MRI scanners are very expensive and they need a very powerful source of electricity.

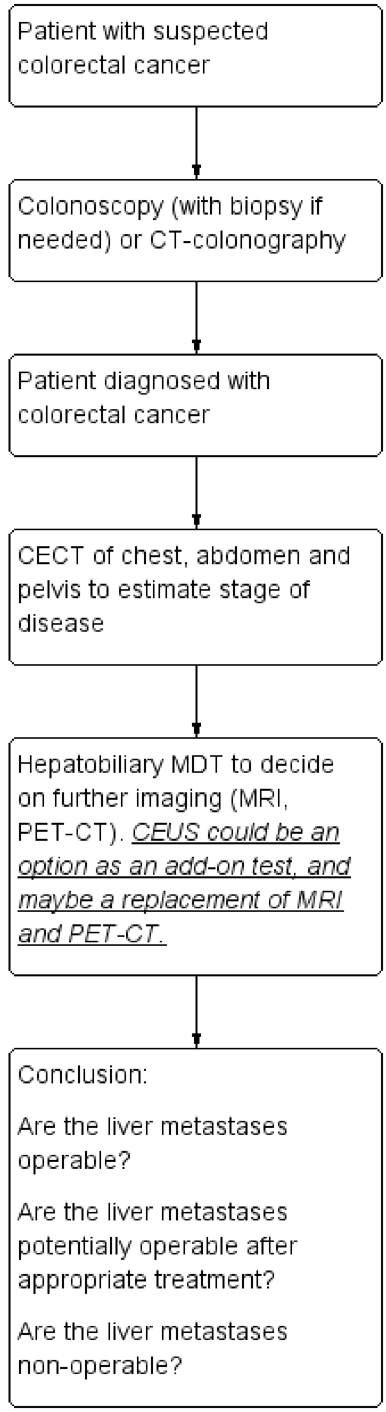
Fluro-18-deoxyglucose positron emission tomography-computed tomography (18F-FDG PET-CT)

A PET-CT scanner is a medical imaging modality based on x-rays and the use of a gamma camera. The CT part works exactly like an ordinary CT scanner, and the gamma camera is a device used to produce images obtained by means of gamma radiation-emitting radio-isotopes, administered as an intravenous infusion. The advantage of 18F-FDG PET-CT is that it provides information on glucose uptake and metabolism of malignant cells in the liver, as well as anatomic alterations like visible liver lesions (Czernin 2010). The disadvantages are the same as for CECT concerning the ionising radiation, but there is no use of contrast agents with iodine. However, the patients are given an intravenous infusion of fluro-18-deoxyglucose (18F-FDG) with an effective dose of approximately 6 to 7 mSv. 18F-FDG PET-CT is a time-consuming examination, and a full scan lasts for around three hours. It has been suggested that hyperglycaemia and diabetes can affect the diagnostic value of 18F-FDG PET-CT, because the cellular uptake of 18F-FDG is adversely affected by elevated plasma glucose levels (Rabkin 2010; Mirpour 2012). However, there was no significant difference in diagnostic accuracy between diabetics and non-diabetics in Rabkin 2010 and Mirpour 2012, and therefore high serum glucose is now no longer considered a contraindication for the conduct of an 18F-FDG PET-CT scan (Rabkin 2010; Mirpour 2012). PET-CT scanners are widely available at the main clinical centres in the western world, but many countries in other parts of the world do not have access to PET-CT scanners. PET-CT-scanners are extremely expensive.

Clinical pathway

CECT is considered in many clinical centres to be the standard imaging technique for detection, diagnosis, and follow-up of people with suspected or verified colorectal liver metastases. The reason for this is that CECT is a very fast and reliable examination, and the images are easily shared among clinicians who have an interest in the examination. The CT images can be manipulated in many different ways, which makes it possible to assess liver lesions from different angles and in both two and three dimensions. This makes CECT a very important tool for the surgeons as well, when they have to assess the possibilities of hepatic resection. If the CECT shows metastatic disease confined to the liver, a multidisciplinary team (MDT) should decide whether further imaging to confirm surgery is suitable for the patient. A CECT or a contrast-enhanced MRI will in most cases conclusively confirm typical colorectal liver metastases. However, if the CECT shows that the person may have extra-hepatic metastases, a MDT should decide whether an 18F-FDG PET-CT of the whole body is appropriate (NICE 2014). Thus, MRI and 18F-FDG PET-CT are used mainly to verify and support findings from CECT (Figure 1). CEUS is currently not a part of the clinical pathway.

Figure 1. Clinical pathway. Inspired of: www.nice.org.uk/guidance/cg131/evidence/full-guideline-183509677



Role of index test(s)

The index tests for this Cochrane diagnostic test accuracy review (DTAR) are CEUS, CECT, MRI, and 18F-FDG PET-CT. The CECT, MRI, and 18F-FDG PET-CT are part of the current clinical pathway, as stated above. To find out whether CEUS has a role in the clinical pathway, we must consider if CEUS should enter the clinical pathway as a replacement test, a triage test, or an add-on test (Bossuyt 2006). A replacement test may be more accurate, less invasive, easier to do, less risky, less uncomfortable for patients, quicker to yield results, technically less challenging, or more easily interpreted than existing tests (Bossuyt 2006). A triage test is used before the existing test(s), and only a particular test result will continue the testing pathway. Triage tests may be less accurate than existing tests, and may not usefully replace them. Triage tests have other advantages, such as simplicity or low cost (Bossuyt 2006). An add-on test may be positioned after the existing pathway if the new test is more accurate, but is otherwise less attractive than existing tests (Bossuyt 2006).

Colorectal cancer may spread to other parts of the body as well as the liver, and CECT can examine hepatic tissue, extra-hepatic tissue and other organs at the same time. For this reason, CECT is a better choice than CEUS as a triage test, because the diagnostic value of CEUS is restricted to the liver. However, CEUS could be a new add-on test if it is as accurate as MRI and 18F-FDG PET-CT in the diagnosis of colorectal liver metastases. From this perspective, CEUS could be a possible replacement for MRI and 18F-FDG PET-CT. From a patient's point of view, it would make sense to choose CEUS as an add-on test to verify and support findings from CECT on colorectal liver metastases, because CEUS does not expose the person to high doses of radiation or contrast agents, which are potentially health-threatening. From an economic context, it may also be valuable to see whether CEUS can compete with MRI and 18F-FDG PET-CT for sensitivity and specificity, because the cost of an ultrasound scanner is much lower than the cost of a MRI scanner or the cost of a PET-CT-scanner.

Rationale

The detection and diagnosis of colorectal liver metastases in people with newly-diagnosed colorectal cancer is important for the staging of the disease. Furthermore, it is important to detect the exact number of liver metastases, their size, their regional distribution, and the volume of the remaining liver in order to determine resectability. The first step in the diagnosis is therefore to determine whether or not liver metastases or extra-hepatic metastases or both are present. CECT is the modality of choice in the current clinical pathway as a triage test for the detection of colorectal liver metastases and extra-hepatic colorectal metastases. MRI and PET-CT are alternative options. The results of the meta-analysis by Niekel 2010 indicate that MRI is the preferred first-line modality for evaluating colorectal liver metastases in people who have not previously undergone therapy.

The second step in the diagnosis is to further investigate all those with a positive triage test result, in order to plan the necessary therapy. It is in this context that CEUS might have a role as an add-on test for people with colorectal metastases confined to the liver, if CEUS is more accurate than CECT, and if CEUS is as accurate as MRI and PET-CT in the diagnosis of colorectal liver metastases (Figure 1).

This Cochrane DTAR aims to provide evidence for the best imaging modality as an add-on test for the diagnosis of liver metastases in people with newly-diagnosed colorectal cancer.

OBJECTIVES

To compare the accuracy of CEUS, CECT, MRI, and 18F-FDG PET-CT for diagnosing liver metastases in people with newly-diagnosed colorectal cancer.

Secondary objectives

Potential sources of heterogeneity

We will investigate the following potential sources of heterogeneity:

- the use of different reference standards.
- different ways of selecting the study populations, e.g. different inclusion and exclusion criteria.
- different locations of the study populations (country, state, region).
- age of participants in the study population.
- sex of participants in the study population.
- differences in clinician skills for the performance of CEUS.

METHODS

Criteria for considering studies for this review

Types of studies

We will include both comparative studies (comparison of two or more index tests against the reference standard), and non-comparative studies (single index test against the reference standard). We will include prospective or retrospective test accuracy cohort

studies, with a cross-sectional design based on a direct comparison of the index tests for diagnosing liver metastases in the same study population of people with newly-diagnosed colorectal cancer, i.e. head-to-head comparisons (e.g. CEUS versus CECT, MRI, or 18F-FDG PET-CT, including any other combination of the modalities). This is the strongest design, especially if it is a fully paired, direct comparison. We will also include studies without fully paired designs, if those studies are randomised accuracy test trials, where study participants are randomly allocated to receive either of the index tests.
We will include cross-sectional studies assessing the accuracy of only one of the index tests (CEUS, CECT, MRI, or 18F-FDG PET-CT) for diagnosing liver metastases in people with newly-diagnosed colorectal cancer, because these studies will show the overall estimate of the accuracy of each index test and allow indirect comparison.

Participants

Adults with newly-diagnosed colorectal cancer. We will not include people with known liver metastases, and people who have already undergone hepatic resection.
Some studies may have restrictions on the inclusion of participants, and we will evaluate these restrictions in every study to reveal the possibilities of bias.

Index tests

The index tests are CEUS, CECT, MRI, and 18F-FDG PET-CT.

Target conditions

The clinical target condition of this review is colorectal liver metastases in people with newly-diagnosed colorectal cancer.

Reference standards

The reference standard should consist of laparotomy including palpation, intraoperative ultrasound (IOUS), and biopsy test results, or a pathological examination of surgically-removed specimens. This approach is considered to be the most accurate. However, this approach is only possible in people with resectable liver metastases. People with non-resectable liver metastases need to be verified by biopsy results. If all of the index tests give negative results in the same study participant, it will not be possible to verify this with the described reference standards, and the participant should be subjected to adequate follow-up for at least three months. At the end of the three months, the participant should once more be evaluated with the same modalities as before, to verify the status of no liver metastases. If the participant is diagnosed with liver metastases after three months, and this is verified by the reference standard, then the first tests should be considered as false negative.
The participants in the same study should all be evaluated by the same reference standard in order to avoid differential verification bias, but because of the facts described above it is necessary to accept more than one reference standard in the same study. However, it is absolutely necessary that all participants in the same category (people with resectable liver metastases, people with non-resectable liver metastases, and people with no liver metastases) are evaluated by the same reference standard.

Search methods for identification of studies

We will discuss the search strategies for relevant studies and establish them in co-operation with the Cochrane Hepato-Biliary Group. We will only perform electronic searches for studies.

Electronic searches

We will conduct electronic searches in The Cochrane Hepato-Biliary Group Controlled Trials Register and The Cochrane Hepato-Biliary Group Diagnostic Test of Accuracy Studies Register (Gluud 2016), the Cochrane Library (Wiley), MEDLINE (PubMed), Embase (OvidSP), and Science Citation Index Expanded (Web of Science) (Royle 2003; De Vet 2008). We will apply no language or document type restrictions. We have provided preliminary search strategies with the expected time spans of the searches in Appendix 1.

Data collection and analysis

We will follow the guidelines provided in the draft *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (De Vet 2008; Reitsma 2013).

Selection of studies

When the search strategies have been established and the search is completed, two authors will independently screen the titles and the abstracts of every study and select the relevant ones individually. We will acquire all the studies selected in this first reading in full text for further assessment. Two authors will each read the full-text studies, and each author will individually identify the relevant studies. We will select the studies in accordance with the inclusion/exclusion criteria as described under the Types of studies section. The authors will then confer and agree upon the studies for inclusion, resolving any disagreements by discussion and consensus. If this approach fails, a third author (the arbiter) will have the final word.

Data extraction and management

Two authors will independently complete a data extraction form for all included studies. We will retrieve the following data:

1. General information: title, journal, year, publication status, and study design.
2. Sample size: number of participants meeting the criteria and total number diagnosed, or scanned, or referred to.
3. Baseline characteristics: baseline diagnosis, age, sex and location (country, state, region).
4. The index test(s).
5. Reference standard.
6. Number of true positive (TP), true negative (TN), false positive (FP), false negative (FN).

Missing data

We will contact study authors for missing or unclear data.

Assessment of methodological quality

Two authors will use QUADAS 2 for the assessment of the methodological quality of the studies (Whiting 2011; Appendix 2). The QUADAS 2 items have been incorporated into the Review Manager 5 software (RevMan 2014). Appendix 2 contains definitions on when to answer yes, no, or unclear to the signalling questions within the QUADAS 2 items, as well as definitions on when the risk of bias is considered high, low, or unclear (Appendix 2). We will resolve any disagreements between the two authors concerning the methodological quality of the studies by consensus. If this approach fails, a third author (the arbiter) will have the final word.

Statistical analysis and data synthesis

We will present each imaging modality in each study as binary data in a 2x2 table. The test results need to be reported as true positive (TP), true negative (TN), false positive (FP), or false negative (FN). We will tabulate and graphically present these values from the selected studies in coupled forest plots (including 95% confidence intervals (CIs)). We will also estimate the positive and negative predictive values, the positive and negative likelihood ratios (LR+ and LR-), and the diagnostic odds ratios with a 95% CI. Furthermore, we will plot the results on a receiver operating characteristic diagram (ROC, sensitivity against 1 - specificity). We will present each of the four index tests in their own ROC space, with the data available from each study. Since we expect a common implicit cut-off between studies, we will use the bivariate model to pool sensitivities and specificities and to estimate the summary operating point (i.e. mean sensitivity and specificity) for each index test. We will perform direct and indirect comparisons by adding the four index tests as covariates to the bivariate model (Reitsma 2005). We will conduct all analyses and plots using Review Manager 5 (RevMan 2014) and Stata (Stata 13).

Investigations of heterogeneity

If we find variability in test accuracy among the studies, we will conduct subgroup analyses by adding covariates to the bivariate model concerning the use of different reference standards, different ways of selecting the study populations (e.g. different inclusion and exclusion criteria), different locations of the study populations (country, state, region), differences between age groups or differences between men and women. If we note variability in test accuracy concerning CEUS among the studies, we will conduct subgroup analyses by adding covariates to the bivariate model, in order to investigate whether the variability is due to differences in clinician skills in the performance of CEUS.

Sensitivity analyses

Sensitivity analyses can be based on the findings of the QUADAS 2. The signalling questions related to the four main domains (participant selection, index test, target condition and reference standard, and flow and timing) will allow allocation of the included studies into three categories: high risk of bias, low risk of bias, and unclear risk of bias. We will rate a study at 'low risk of bias' if the answers to all the signalling questions are "yes". We plan to conduct a sensitivity analysis by including only studies at low risk of bias.

ACKNOWLEDGEMENTS

Comments were received from Mirella Fraquelli, Italy and Christian Gluud, Denmark.

We thank Sarah Louise Klingenberg, Denmark, for drafting preliminary search strategies.

Contact Editors at the Cochrane Hepato-Biliary Group Diagnostic Test Accuracy Reviews Editorial Team, Denmark: Agostino Colli, Italy and Giovanni Casazza, Italy.

Cochrane Review Group funding acknowledgement: The Danish State is the largest single funder of The Cochrane Hepato-Biliary Group through its investment in The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Denmark. Disclaimer: The views and opinions expressed in this review are those of the authors and do not necessarily reflect those of the Danish State or The Copenhagen Trial Unit.

REFERENCES

Additional references

Bossuyt 2006
Bossuyt PM, Irwig L, Craig J, Glasziou P. Comparative accuracy: assessing new tests against existing diagnostic pathways. *BMJ (Clinical Research Ed.)* 2006;**332**:1089–92.

Cirocchi 2012
Cirocchi R, Trastulli S, Boselli C, Montedori A, Cavaliere D, Parisi A, et al. Radiofrequency ablation in the treatment of liver metastases from colorectal cancer. *Cochrane Database of Systematic Reviews* 2012, Issue 6. [DOI: 10.1002/14651858.CD006317.pub3; CD006317]

Crawford 2012
Crawford F, McIntosh H, Dozier M, Brush J, Glanville JM, Dunlop MG, et al. FDG PET-CT imaging for pre operative staging in patients with colorectal cancer. *Cochrane Database of Systematic Reviews* 2012, Issue 2. [DOI: 10.1002/14651858.CD009630; CD009630]

Czernin 2010
Czernin J, Benz MR, Allen-Auerbach MS. PET/CT imaging: The incremental value of assessing the glucose metabolic phenotype and the structure of cancers in a single examination. *European Journal of Radiology* 2010;**73**(3):470–80. [PUBMED: 20097498]

De Vet 2008
de Vet HCW, Eisinga A, Riphagen II, Aertgeerts B, Pewsner D. Chapter 7: Searching for studies. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 [updated September 2008]*. The Cochrane Collaboration, 2008.

Gluud 2016
Gluud C, Nikolova D, Klingenberg SL. Cochrane Hepato-Biliary Group. About Cochrane (Cochrane Review Groups (CRGs)) 2016, Issue 4. Art. No.: LIVER.

Kanas 2012
Kanas GP, Taylor A, Primrose JN, Langeberg WJ, Kelsh MA, Mowat FS, et al. Survival after liver resection in metastatic colorectal cancer: review and meta-analysis of prognostic factors. *Clinical Epidemiology* 2012;**4**:283–301. [PUBMED: 23152705]

Lim 2014
Lim MC, Tan CH, Cai J, Zheng J, Kow AW. CT volumetry of the liver: where does it stand in clinical practice?. *Clinical Radiology* 2014;**69**(9):887–95. [PUBMED: 24824973]

Mirpour 2012
Mirpour S, Meteesatien P, Khandani AH. Does hyperglycemia affect the diagnostic value of 18F-FDG PET/CT?. *Revista Española de Medicina Nuclear e Imagen Molecular* 2012;**31**(2):71–7. [PUBMED: 22088805]

NCHS 2010
National Center for Health Statistics. Health, United States, 2010: With Special Feature on Death and Dying. Hyattsville, MD. 2011. www.cdc.gov/nchs/data/hus/hus10.pdf (accessed 7 June 2016).

NICE 2014
National Institute for Health and Care Ecellence (NICE). Colorectal cancer: diagnosis and management. NICE guidelines [CG131]. www.nice.org.uk/guidance/cg131 2014 (accessed 7 June 2016).

Niekel 2010
Niekel MC, Bipat S, Stoker J. Diagnostic imaging of colorectal liver metastases with CT, MR imaging, FDG PET, and/or FDG PET/CT: a meta-analysis of prospective studies including patients who have not previously undergone treatment. *Radiology* 2010;**257**(3):674–84. [PUBMED: 20829538]

Rabkin 2010
Rabkin Z, Israel O, Keidar Z. Do hyperglycemia and diabetes affect the incidence of false-negative 18F-FDG PET/CT studies in patients evaluated for infection or inflammation and cancer? A comparative analysis. *Journal of Nuclear Medicine* 2010;**51**(7):1015–20. [PUBMED: 20554733]

Reitsma 2005
Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982–90. [PUBMED: 16168343]

Reitsma 2013
Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ, editor(s). Chapter 9: Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C, editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0, The Cochrane Collaboration, 2013 Available from: srdta.cochrane.org/.

RevMan 2014 [Computer program]
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Royle 2003
Royle P, Milne R. Literature searching for randomized controlled trials used in Cochrane reviews: rapid versus exhaustive searches. *International Journal of Technology Assessment in Health Care* 2003;**19**(4):591–603.

Siegel 2015
Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA: a Cancer Journal for Clinicians* 2015;**65**(1):5–29. [PUBMED: 25559415]

Solbiati 2003
Solbiati L, Martegani A, Leen E, Correas JM, Burns PN, Becker D. *Contrast-Enhanced Ultrasound of Liver Diseases*. Milan: Springer-Verlag Italy, 2003.

Stata 13 [Computer program]
StataCorp LP. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP, 2013.

Vatandoust 2015
Vatandoust S, Price TJ, Karapetis CS. Colorectal cancer: metastases to a single organ. *World Journal of Gastroenterology* 2015;**21**(41):11767–76. [PUBMED: 26557001]

Westbrook 2011
Westbrook C, Roth CK, Talbot J. *MRI in Practice*. 4th Edition. Wiley-Blackwell, 2011.

Whiting 2011
Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529–36. [PUBMED: 22007046]

* *Indicates the major publication for the study*

APPENDICES

Appendix 1. Preliminary search strategies

Database	Time span	Search strategy
The Cochrane Hepato-Biliary Group Controlled Trials Register	Date will be given at review stage.	((computed tomograph* or CT or CECT or MDCT or MSCT or magnetic resonance imaging or MRI or emission tomography or PET) OR (ultrasound or ultrasonograph* or US or CEUS)) AND ((liver or hepat*) near (metasta* or secondar* or spread or advanced)) AND ((colorectal or rectal or colon) near/3 (cancer or carcinom* or neoplasm* or tumo?r*))
The Cochrane Hepato-Biliary Group Diagnostic Test of Accuracy Studies Register	Date will be given at review stage.	((computed tomograph* or CT or CECT or MDCT or MSCT or magnetic resonance imaging or MRI or emission tomography or PET) OR (ultrasound or ultrasonograph* or US or CEUS)) AND ((liver or hepat*) near (metasta* or secondar* or spread or advanced)) AND ((colorectal or rectal or colon) near/3 (cancer or carcinom* or neoplasm* or tumo?r*))
The Cochrane Library (Wiley)	Latest issue	#1 MeSH descriptor: [Tomography, Emission-Computed] explode all trees #2 MeSH descriptor: [Tomography, X-Ray Computed] explode all trees #3 MeSH descriptor: [Magnetic Resonance Imaging] explode all trees

(*Continued*)

		#4 (computed tomograph* or CT or CECT or MDCT or MSCT or magnetic resonance imaging or MRI or emission tomography or PET) #5 #1 or #2 or #3 or #4 #6 MeSH descriptor: [Ultrasonography] explode all trees #7 ultrasound or ultrasonograph* or US or CEUS #8 #6 or #7 #9 MeSH descriptor: [Liver Neoplasms] explode all trees #10 (liver or hepat*) near (metasta* or secondar* or spread or advanced) #11 #9 or #10 #12 MeSH descriptor: [Colorectal Neoplasms] explode all trees #13 (colorectal or rectal or colon) near/3 (cancer or carcinom* or neoplasm* or tumo?r*) #14 #12 or #13 #15 (#5 or #8) and #11 and #14
MEDLINE (OvidSP)	1946 to the date of search.	1. exp Tomography, Emission-Computed/ 2. exp Tomography, X-Ray Computed/ 3. exp Magnetic Resonance Imaging/ 4. (computed tomograph* or CT or CECT or MDCT or MSCT or magnetic resonance imaging or MRI or emission tomography or PET). mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 5. 1 or 2 or 3 or 4 6. exp Ultrasonography/ 7. (ultrasound or ultrasonograph* or US or CEUS) .mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 8. 6 or 7 9. exp Liver Neoplasms/ 10. ((liver or hepat*) adj (metasta* or secondar* or spread or advanced)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 11. 9 or 10

(*Continued*)

		12. exp Colorectal Neoplasms/ 13. ((colorectal or rectal or colon) adj3 (cancer or carcinom* or neoplasm* or tumo?r*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier] 14. 12 or 13 15. (5 or 8) and 11 and 14
Embase (OvidSP)	1974 to the date of search.	1. exp computer assisted tomography/ 2. exp positron emission tomography/ 3. exp nuclear magnetic resonance imaging/ 4. (computed tomograph* or CT or CECT or MDCT or MSCT or magnetic resonance imaging or MRI or emission tomography or PET). mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] 5. 1 or 2 or 3 or 4 6. exp echography/ 7. exp ultrasound/ 8. (ultrasound or ultrasonograph* or US or CEUS) .mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] 9. 6 or 7 or 8 10. exp liver metastasis/ 11. ((liver or hepat*) adj (metasta* or secondar* or spread or advanced)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] 12. 10 or 11 13. exp colorectal cancer/ 14. ((colorectal or rectal or colon) adj3 (cancer or carcinom* or neoplasm* or tumo?r*)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] 15. 13 or 14 16. (5 or 9) and 12 and 15
Science Citation Index Expanded (Web of Science)	1900 to the date of search.	#5 (#1 or #2) AND #3 AND #4 #4 TS=((colorectal or rectal or colon) near/3 (cancer or carcinom* or neoplasm* or tumo?r*)) #3 TS=((liver or hepat*) near (metasta* or secondar* or spread or advanced)) #2 TS=(ultrasound or ultrasonograph* or US or

(*Continued*)

		CEUS) #1 TS=(computed tomograph* or CT or CECT or MDCT or MSCT or magnetic resonance imaging or MRI or emission tomography or PET)

Appendix 2. QUADAS 2 items

DOMAIN	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING
Description	**Describe methods of patient selection: Describe included participants (prior testing, presentation, intended use of index test and setting):**	**Describe the index test and how it was conducted and interpreted:**	**Describe the reference standard and how it was conducted and interpreted:**	**Describe any participants who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard:**
Signalling questions (yes/no/unclear)	**Was a consecutive or random sample of participants enrolled?** Yes: a consecutive or random sample of participants with newly diagnosed colorectal cancer were enrolled in the study No: selection of participants was reported. Unclear: insufficient data were reported to permit a judgment	**Were the index test results interpreted without knowledge of the results of the reference standard?** Yes: the index tests results were interpreted blinded to the results of the reference standard No: the index tests results were not interpreted blinded to the results of the reference standard Unclear: insufficient data were reported to permit a judgment	**Is the reference standard likely to correctly classify the target condition?** Yes: If all patients have undergone the reference standard as described No: If not all participants have undergone the reference standard as described Unclear: insufficient data were reported to permit a judgment	**Was there an appropriate interval between index test(s) and reference standard?** Yes: the period of time for the index tests and the reference standard to be executed was shorter than or equal to three months No: the period of time for the index tests and the reference standard to be executed was longer than three months Unclear: insufficient data were reported to permit a judgment
	Was a case-control design avoided? Yes: case-control design was avoided.	**If a threshold was used, was it pre-specified?** Yes: a threshold was used, and the threshold was	**Were the reference standard results interpreted without knowledge of the results of**	**Did all participants**receive the reference standard?** Yes: all participants re-

(*Continued*)

	No: case-control design was not avoided. Unclear: insufficient information was reported to permit a judgment	pre-specified Or: A threshold was not used. No: a threshold was used, and the threshold was not pre-specified Unclear: insufficient data were reported to permit a judgment	**the index test?** Yes: the reference standard results were interpreted blinded to the results of the index tests No: the reference standard results were not interpreted blinded to the results of the index tests Unclear: insufficient data were reported	ceived the reference standard (see description under the headline "Reference standards".) No: not all participants received the reference standard (see description under the headline "Reference standards".) Unclear: insufficient data were reported
	Did the study avoid inappropriate exclusions? Yes: the study avoided exclusions of participants who were difficult to diagnose, e.g. not clearly positive or negative test results due to sub optimal examinations No: the study excluded participants who were difficult to diagnose Unclear: insufficient data were reported to permit a judgment			**Did all participants receive the same reference standard?** Yes: all participants received the same reference standard (see description under the headline "Reference standards".) No: not all participants received the same reference standard (see description under the headline "Reference standards".) Unclear: insufficient data were reported to permit a judgment
				Were all participants included in the analysis? Yes: all participants meeting the selection criteria (selected participants) were included in the analysis, or data on all the selected participants were available so that a 2 x 2 table including all selected participants could be constructed No: not all participants meeting the selection criteria (selected partici-

(*Continued*)

				pants) were included in the analysis, and data on all the selected participants were not available so that a 2 x 2 table could be constructed using data on all selected participants Unclear: insufficient data were reported to permit a judgment
Risk of bias: High/low/unclear	**Could the selection of participants have introduced bias?** High risk of bias: if at least one of the answers to the signalling questions on the selection of participants are "no" Low risk of bias: if all the answers to the signalling questions on the selection of participants are "yes" Unclear risk of bias: if the answers to the signalling questions on the selection of participants are either "unclear" or any combination of "unclear" with "yes" and/or "no"	**Could the conduct or interpretation of the index test have introduced bias?** High risk of bias: if at least one of the answers to the signalling questions on the conduct or interpretation of the index tests are "no" Low risk of bias: if all the answers to the signalling questions on the conduct or interpretation of the index tests are "yes" Unclear risk of bias: if the answers to the signalling questions on the conduct or interpretation of the index tests are either "unclear" or any combination of "unclear" with "yes" and/or "no"	**Could the reference standard, its conduct, or its interpretation have introduced bias?** High risk of bias: if at least one of the answers to the signalling questions on the conduct or interpretation of the reference standard are "no" Low risk of bias: if all the answers to the signalling questions on the conduct or interpretation of the reference standard are "yes" Unclear risk of bias: if the answers to the signalling questions on the conduct or interpretation of the reference standard are either "unclear" or any combination of "unclear" with "yes" and/or "no"	**Could the patient flow have introduced bias?** High risk of bias: if at least one of the answers to the signalling questions on flow and timing are "no" Low risk of bias: if all the answers to the signalling questions on flow and timing are "yes" Unclear risk of bias: if the answers to the signalling questions on flow and timing are either "unclear" or any combination of "unclear" with "yes" and/or "no"
Concerns regarding applicability: High/low/unclear	**Are there concerns that the included participants do not match the review question?** High concern: if the included participants do not match the description under the headline "Participants". Low concern: if the in-	**Are there concerns that the index test, its conduct, or interpretation differ from the review question?** High concern: if the conduct or interpretation of the index tests do not match the description under the headline	**Are there concerns that the target condition as defined by the reference standard does not match the review question?** High concern: if the participant are diagnosed with any other cancer disease than colorectal	

(*Continued*)

	cluded participants match the description under the headline "Participants". Unclear: If it is unclear whether the included participants match the description under the headline "Participants" or not.	"Index test(s)". Low concern: if the conduct and interpretation of the index tests match the description under the headline "Index test(s)". Unclear: If it is unclear whether the conduct and/or interpretation of the index tests match the description under the headline " Index test(s)" or not.	cancer Low concern: if the participant are diagnosed with colorectal cancer Unclear: If it is unclear wether the participant are diagnosed with colorectal cancer or not	

CONTRIBUTIONS OF AUTHORS

Martin Lund: content expertise, review expertise, proposal and writing of the protocol, selection of studies, assessment of selected studies, writing of the review.

Thomas A Bjerre: clinical expertise, selection of studies, assessment of selected studies.

Henning Grønbæk: clinical expertise, arbiter.

Frank V Mortensen: clinical expertise.

Per Kragh Andersen: statistical expertise, review expertise, statistical calculations.

All authors approved the current protocol content.

DECLARATIONS OF INTEREST

Martin Lund: none known.

Thomas A Bjerre: none known.

Henning Grønbæk: research grants from Novartis, Ipsen, Abbie, Intercept. Principal investigator for studies, sponsored by Ipsen, Novartis, Intercept. However, none of the listed declared conflicts is relevant to this review.

Frank V Mortensen: none known.

Per Kragh Andersen: none known.